Wavelet-based denoising for 3D OCT images

Vladimir Zlokolica[1], Ljubomir Jovanov[1], Aleksandra Pižurica[1], Paul De Keyser[2], Frans Dhaenens[2] and Wilfried Philips[1]

[1] Ghent University, Dept. of Telecommunications and Information Processing (TELIN), IPI, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium;
[2] Agfa-Gevaert N.V., Septestraat 27, B-2640 Mortsel, Belgium

ABSTRACT

Optical coherence tomography produces high resolution medical images based on spatial and temporal coherence of the optical waves backscattered from the scanned tissue. However, the same coherence introduces speckle noise as well; this degrades the quality of acquired images.

In this paper we propose a technique for noise reduction of 3D OCT images, where the 3D volume is considered as a sequence of 2D images, i.e., 2D slices in depth-lateral projection plane. In the proposed method we first perform recursive temporal filtering through the estimated motion trajectory between the 2D slices using noise-robust motion estimation/compensation scheme previously proposed for video denoising. The temporal filtering scheme reduces the noise level and adapts the motion compensation on it. Subsequently, we apply a spatial filter for speckle reduction in order to remove the remainder of noise in the 2D slices. In this scheme the spatial (2D) speckle-nature of noise in OCT is modeled and used for spatially adaptive denoising. Both the temporal and the spatial filter are wavelet-based techniques, where for the temporal filter two resolution scales are used and for the spatial one four resolution scales.

The evaluation of the proposed denoising approach is done on demodulated 3D OCT images on different sources and of different resolution. For optimizing the parameters for best denoising performance fantom OCT images were used. The denoising performance of the proposed method was measured in terms of SNR, edge sharpness preservation and contrast-to-noise ratio. A comparison was made to the state-of-the-art methods for noise reduction in 2D OCT images, where the proposed approach showed to be advantageous in terms of both objective and subjective quality measures.

Keywords: 3D OCT, wavelets, denoising

1. INTRODUCTION

Optical Coherence Tomography performs high resolution, cross-sectional tomographic imaging of the internal micro-structure in materials and biological systems by measuring backscattered or back-reflected light [1]. OCT is capable of real-time in vivo imaging with multiple frames per second with little image degradation due to motion of the specimen [2].

Although OCT produces high resolution images, it also introduces significant amount of noise. The noise complicates many post-processing tasks such as detection of important features and degrades visual quality in general. The acquired noise is usually spatially correlated and signal dependent, often referred to as "speckle" noise.

Speckle-noise reduction of OCT images aims at suppressing signal-degrading speckle and accentuating noise-free signal. Numerous methods have been developed for speckle reduction in OCT images; most common are median filtering [3], homomorphic Wiener filtering [4], multi-resolution based techniques [5] and adaptive smoothing [6].

In this work, we propose a filtering scheme for OCT images, which combines our video denoising algorithm from [7] and speckle noise reduction from [8]. Specifically, we apply first spatio-temporal filtering

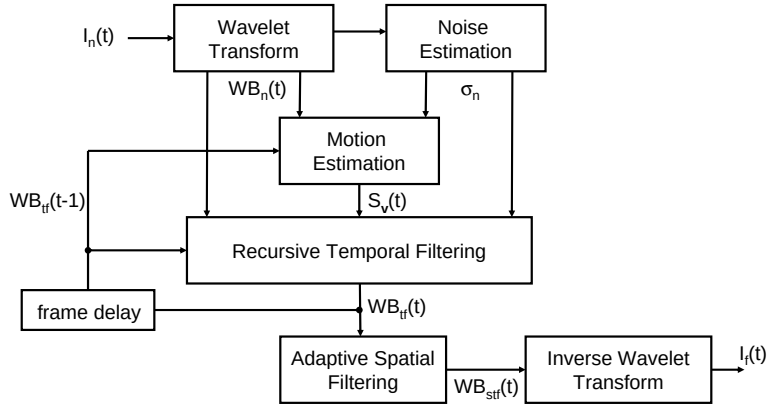

Figure 1. The general framework description of the proposed algorithm: $I_n(t)$ - input 2D OCT noisy frame, $I_f(t)$ - spatio-temporally filtered frame, $WB_n(t)$ - noisy wavelet band, $WB_{tf}(t)$ temporally filtered wavelet band; $WB_{stf}(t)$ spatio-temporally filtered wavelet band; σ_n - standard deviation of Gaussian noise; t - temporal coordinate; $S_{\mathbf{v}}(t)$ - set of estimated motion vectors $\mathbf{v}$ for frame t.

using the method of [7] and for the suppression of the remaining noise we propose a simple three-dimensional extension of the method of [8].

In Section 2 we present the proposed spatio-temporal scheme for noise reduction. First we describe the motion compensated filtering scheme from [7] (Section 2.1) and then we present the proposed extension of the speckle filter from [8] (Section 2.2). In Section 3 we present experimental results and in Section 4 we conclude the paper and give hints for future work.

2. THE PROPOSED ALGORITHM FOR WAVELET-BASED DENOISING 3D OCT IMAGES

In our denoising approach we treat three-dimensional (3D) OCT images as sequences of two-dimensional (2D) slices (frames). Since the processed 3D OCT image sequences were obtained by consecutive acquiring (scanning) 2D OCT images in direction perpendicular to the acquired image plane at different, fixed distances (and at different time instances), we treat the 3D OCT volume as a sequence of 2D OCT image frames and we treat changes from one frame to another as motion (in the "temporal" direction). Consequently, we apply a causal recursive temporal filter to each OCT slice. The denoising of each slice is hence influenced by the previously processed OCT slices.

A general description of the proposed denoising algorithm is presented in Fig. 1. Three important steps are:

- *Motion estimation*: The proposed approach estimates a *single* motion vector field for all wavelet bands but *different* motion vector reliabilities in each band. In the motion estimation, we make use of spatial image discontinuities, represented by wavelet coefficients, in combination with spatio-temporal correlations of the motion vectors, within a recursive scheme.
- *Motion compensation*: Temporal filtering in each wavelet band is recursively performed along the estimated motion trajectories. The amount of filtering is tuned to the estimated reliability of the corresponding motion vectors within a certain block of wavelet coefficients.
- *The adaptive spatial filtering scheme* suppresses the remaining (non-stationary) noise. It evaluates distributions of the wavelet coefficients which represent mainly speckle noise on the one hand and useful signal corrupted by speckle on the other. Analytic models for evaluating these distributions is given in [8] for 2D images with their parameters automatically computed from a given 2D image.

In the denoising scheme, motion estimation and compensation were proposed in [7] for video sequences with white Gaussian noise; we apply here the same framework on a new application: OCT image sequences, as a part of the denoising framework, in combination with [8].

The proposed method uses a non-decimated wavelet transform implemented with spline wavelets [9], which results in two detail images and one approximation (low-pass) image at each scale. We apply a two-dimensional (spatial) wavelet transform to each 2D OCT frame and denote *wavelet bands* of this spatial wavelet transform by $WB = LL, DX, DY$ for the low-pass (approximation), horizontal and vertical orientation bands, respectively. We use a subscript to denote the noisy or denoised band as follows: WB_n - noisy band, WB_{tf} - temporally filtered and WB_{stf} - spatio-temporally filtered band. Additionally, we denote the spatial position as $\mathbf{r} = (x, y)$ and frame index (time) as t. The decomposition level is denoted by a superscript (l), where $l = 1, \ldots, N$ (1 denotes the finest scale and N the coarsest).

As can be seen from Fig. 1, the noisy input frame $I_n(\mathbf{r}, t)$ is first decomposed into wavelet bands $WB_n^{(l)}(\mathbf{r}, t)$. Using the noisy wavelet bands $WB_n^{(l)}(t)$ from the current frame and temporally filtered wavelet bands $WB_{tf}^{(l)}(t-1)$ from the previous time recursion, and taking into account the estimated standard deviation (σ_n) of the Gaussian noise*, we perform motion estimation. The estimation of the (approximated) Gaussian noise level is performed using a spatial-gradient approach [10], based on evaluating the distribution of spatial gradient magnitudes. Subsequently, we apply a recursive temporal filter (Subsection 2.1.1) on the noisy wavelet bands $WB_n^{(l)}(t)$ along the estimated motion trajectory, using the corresponding $WB_{tf}^{(l)}(t-1)$ band. The temporally filtered wavelet bands $WB_{tf}^{(l)}(t)$ are further subjected to spatial filtering. Finally, the inverse wavelet transform yields the denoised video sequence $I_f(t)$.

2.1. Inter-based Denoising of 2D slices in z-direction

Classical single resolution motion estimation and compensation techniques use block-matching [11, 12]. This assumes that motion is translational and locally stationary. The motion vector $\mathbf{v}$ is determined by minimizing a certain *cost* function in a confined search area:

$$\mathbf{v}(s) = \arg\min_{\mathbf{v}}[\text{cost}(s, \mathbf{v})] \tag{1}$$

where $\text{cost}(s, \mathbf{v})$ is a linear or nonlinear function of the pixel values from the s-th block of the current frame and the pixel values from the corresponding block in the previous frame, displaced by the motion vector $\mathbf{v}$. The most common cost function is the mean absolute difference (*MAD*).

In [7], we introduced a novel block-based motion estimation approach which employs a three-step search [11–13], within the wavelet domain, operating in a spatio-temporal recursive manner. In the first step the initial motion vector is either equal to the previously computed motion vector from the spatio-temporal neighborhood or it is set to a zero vector. Subsequently, the chosen initial motion vector is refined in the three steps. The proposed motion estimation approach produces a single motion vector field which is used for the adaptive temporal filtering for all wavelet bands.

In the proposed method, the refinement of the initial motion vector is based on motion-matching image discontinuities (represented by groups of significant wavelet coefficients) and defined reliabilities *per orientation* (2). Specifically, the motion matching value of the image discontinuities, of a specific orientation (horizontal or vertical), is determined as a weighted sum of MAD's (3) for the tested motion vector and corresponding wavelet bands of the same orientation and different resolution scales.

Further on, in order to properly coordinate our motion estimation process we introduce the concept of reliability per orientation for the corresponding initial motion vector that is to be refined. The idea of using motion reliability is as follows [7]: we assume that the motion estimation is most reliable on significant image discontinuities (producing largest wavelet coefficient magnitudes) and the least reliable in uniform

*We make here an approximation that the noise is white Gaussian and disregard its spatial correlation.

areas (characterized by none-significant wavelet coefficient values). Consequently, we aim at estimating motion vectors in highly reliable areas, by performing refinement of the initially estimated motion vector; in less reliable areas we rather rely on (already obtained) reliable spatio-temporally neighboring motion vectors and prevent further refinements (changes) of the chosen initial motion vector.

We define the horizontal θ_H and vertical θ_V "per orientation" reliabilities of the motion vector $\mathbf{v}$, as follows:

$$\begin{aligned}\theta_H(s,t,\sigma_n,\mathbf{v}) &= \frac{\sigma_n}{1+\sum_{l=1}^{N} d_l MAD_{DY}^{(l)}(s,t,\mathbf{v})}\\ \theta_V(s,t,\sigma_n,\mathbf{v}) &= \frac{\sigma_n}{1+\sum_{l=1}^{N} d_l MAD_{DX}^{(l)}(s,t,\mathbf{v})}\end{aligned} \tag{2}$$

with parameters d_l $(l = 1, ..., N)$ denoting the weights associated with the corresponding wavelet decomposition scale l, where $d_1 = d_2 = \ldots = d_N$ and $\sum_{l=1}^{N} d_l = 1$.

The Mean Absolute Difference (MAD) for each block s in the wavelet band $WB^{(l)}(\mathbf{r}, t)$ is defined as follows:

$$MAD_{WB}^{(l)}(s,t,\mathbf{v}) = \frac{1}{N}\sum_{\mathbf{r}\in B_s} |WB_n^{(l)}(\mathbf{r},t) - WB_{tf}^{(l)}(\mathbf{r}-\mathbf{v},t-1)| \tag{3}$$

where s denotes the index of a block within the current frame, t the current frame index and $\mathbf{v}$ a motion vector. B_s represents a set of $N = N_x \times N_y$ spatial positions belonging to the given block s, where $N_x = 8$ and $N_y = 8$ represent the number of rows and columns in the block, respectively.

Additionally, after final motion vector $\mathbf{v}$ has been estimated for particular block s, we determine the reliability of the finally estimated motion vectors *per wavelet band* to determine the appropriate *amount* of temporal filtering (Subsection 2.1.1), along the estimated motion trajectory. The "per wavelet band" reliability for the $WB^{(l)}$) of the estimated motion vector $\mathbf{v}$, for temporal filtering, is defined as follows [7]:

$$\theta_{WB}^{(l)}(s,t,\sigma_n,\mathbf{v}) = \frac{\sigma_n}{1+MAD_{WB}^{(l)}(s,t,\mathbf{v})}. \tag{4}$$

In our motion estimation approach the search area in each step of the algorithm is confined to a $(N_x/2^{j-1}) \times (N_y/2^{j-1})$ block where $j = 1, 2, 3$ for the first, second and third step, respectively. For each step j of the algorithm, we define initial motion vectors $\mathbf{v}_i^{(j)}$ and motion vector corrections $\mathbf{v}_c^{(j)}$ where $\mathbf{v}_i^{(j)}$ is fixed and $\mathbf{v}_c^{(j)}$ varies. The vectors $\mathbf{v}_i^{(j)} + \mathbf{v}_c^{(j)}$ are then tested in order to find the best matching vector $\mathbf{v}_b^{(j)}$ for step j. The best matching vector is determined by minimizing a cost function (which we define later), as follows:

$$\mathbf{v}_b^{(j)}(s,t) = \arg\min_{\mathbf{v}_c^{(j)}\in S_j} \text{cost}(s,t,\vartheta_H,\vartheta_V,\mathbf{v}_i^{(j)},\mathbf{v}_c^{(j)}) \tag{5}$$

where S_j represents the set of allowed $\mathbf{v}_c^{(j)}$s. The best matching motion vector $\mathbf{v}_b^{(j)}$ at step three $(j = 3)$ is defined as the final estimated motion vector $\mathbf{v}_{bf}$. The initial motion vector at each step $j > 1$ is the best estimated motion vector from the preceding step, i.e., $\mathbf{v}_i^{(j)} = \mathbf{v}_b^{(j-1)}$. In the first step $(j = 1)$ the initial motion vector is the best matching motion vector among the tested motion vectors from a spatio-temporal neighborhood. Specifically, we define the initial motion vector at step 1, $\mathbf{v}_i^{(1)}$ as the prior initial motion vector candidate $\mathbf{v}_{pi}$ that minimizes:

$$\mathbf{v}_i^{(1)}(s,t) = \arg\min_{\mathbf{v}_{pi}\in U} \Big(MAD_{LL}^{(N)}(s,t,\mathbf{v}_{pi}) + P(\mathbf{v}_{pi})\Big) \tag{6}$$

where $\mathbf{v}_{pi}$ denotes the prior initial motion vector candidate and $P(\mathbf{v}_{pi})$ represents a penalty for the corresponding vector $\mathbf{v}_{pi}$, with which we introduce prior knowledge. If the penalty $P(\mathbf{v}_{pi})$ is smaller, the initial

motion vector candidate $\mathbf{v}_{pi}$ is more likely to be chosen as the initial motion vector $\mathbf{v}_i^{(1)}$ for the first step of the motion estimation approach.

The prior initial motion vector candidates $\mathbf{v}_{pi}$ belong to set $U = \{\mathbf{0}, \mathbf{s}, \mathbf{s}', \mathbf{t}, \mathbf{t}'\}$. This candidate set includes the zero motion vector ($\mathbf{0}$) and the motion vectors from two neighboring blocks within the current frame ($\mathbf{s}, \mathbf{s}'$) and from two neighboring blocks in the previous frame ($\mathbf{t}, \mathbf{t}'$). The zero motion vector ($\mathbf{0}$) is used for a re-initialization of the motion vector search (estimation) and the spatio-temporal neighboring vectors ($\mathbf{s}, \mathbf{s}', \mathbf{t}, \mathbf{t}'$) are used to enable spatio-temporal recursiveness in the motion estimation approach.

By assigning the smallest penalty to $\mathbf{v}_{pi} = \mathbf{0}$, we increase the sensitivity of the motion estimation to sudden scene changes or the appearance of small image parts (this concerns the accuracy of the motion estimation). Hence, the penalty for $\mathbf{v}_{pi}$ being equal to either of the four spatio-temporal neighboring vectors should be sufficiently large to re-initialize the motion vector search in case of sudden scene changes. However, the penalty value should not be too big either in order to enable spatio-temporal recursiveness in the motion estimation and consequently enforce smoothness (consistency) of the motion vector field. In our experiments, we use $P(\mathbf{0}) = 0$ and $P(\mathbf{v}_{pi}) = 2.5$; this constant was experimentally optimized in terms of maximal motion vector consistency and accuracy.

In step 1, we consider the candidate motion vector corrections $\mathbf{v}_c^{(1)}$ with horizontal component $v_{c_x}^{(1)}$ and vertical component $v_{c_y}^{(1)}$ in the set $\{-8, -4, 0, 4, 8\}$. They are added to the $\mathbf{v}_i^{(1)}$ and tested in order to find the best matching motion vector $\mathbf{v}_b^{(1)}$. In the second step, $\mathbf{v}_i^{(2)} = \mathbf{v}_b^{(1)}$ and the candidate motion vector corrections $\mathbf{v}_c^{(2)}$ can have horizontal $v_{c_x}^{(2)}$ and vertical $v_{c_y}^{(2)}$ components within the set $\{-4, -2, 0, 2, 4\}$. Finally, in the third step, $\mathbf{v}_i^{(3)} = \mathbf{v}_b^{(2)}$ and the candidate motion vector correction $\mathbf{v}_c^{(3)}$ components $v_{c_x}^{(3)}$ and $v_{c_y}^{(3)}$ belong to the set $\{-2, -1, 0, 1, 2\}$.

We define a novel cost function for motion estimation, consisting of horizontal and vertical components, where each component is weighted by the estimated reliability measures with respect to the corresponding initial motion vector component:

$$\begin{aligned} \text{cost}(s, t, \vartheta_H, \vartheta_V, \mathbf{v}_i, \mathbf{v}_c) &= k_x(s, t, \vartheta_H, v_{c_x})\text{cost}_x(s, t, \mathbf{v}_i + \mathbf{v}_c) \\ &\quad + k_y(s, t, \vartheta_V, v_{c_y})\text{cost}_y(s, t, \mathbf{v}_i + \mathbf{v}_c), \end{aligned} \tag{7}$$

where $\text{cost}_x(s, t, \mathbf{v})$ and $\text{cost}_y(s, t, \mathbf{v})$ are separate cost functions for horizontal and vertical motion, respectively, defined as:

$$\begin{aligned} \text{cost}_x(s, t, \mathbf{v}) &= MAD_{LL}^{(N)}(s, t, \mathbf{v}) + \textstyle\sum_{l=1}^{N} MAD_{DY}^{(l)}(s, t, \mathbf{v}) \\ \text{cost}_y(s, t, \mathbf{v}) &= MAD_{LL}^{(N)}(s, t, \mathbf{v}) + \textstyle\sum_{l=1}^{N} MAD_{DX}^{(l)}(s, t, \mathbf{v}). \end{aligned} \tag{8}$$

The multiplicative penalties k_x and k_y in the cost function (7) are defined in terms of the motion reliabilities, as follows:

$$\begin{aligned} k_x(s, t, \vartheta_H, v_{c_x}^{(j)}) &= C_1 + C_2 \frac{|v_{c_x}^{(j)}|}{2^{N-j}} \vartheta_H^2, \\ k_y(s, t, \vartheta_V, v_{c_y}^{(j)}) &= C_1 + C_2 \frac{|v_{c_y}^{(j)}|}{2^{N-j}} \vartheta_V^2, \end{aligned} \tag{9}$$

where the constants C_1 and C_2 are optimized in order to obtain a noise robust and smooth motion vector field. We have experimentally found the following optimal parameter values: $C_1 = 1$, $C_2 = 1.45$. The values of the constants C_1 and C_2 are fixed in all three steps and the correction motion vector components ($|v_{c_x}^{(j)}|$ and $|v_{c_x}^{(j)}|$) are normalized with their maximum amplitudes (2^{N-j}) for j step of the proposed algorithm.

2.1.1. Recursive Temporal Filtering

The proposed wavelet domain recursive temporal filtering (RTF) scheme filters a video sequence along the estimated motion trajectories and adapts the amount of smoothing to the estimated reliability of the motion vectors. Specifically, recursive adaptive temporal filtering is performed separately in each noisy (non-processed) wavelet band $WB_n^{(l)}(\mathbf{r},t)$ as follows:

$$WB_{tf}^{(l)}(\mathbf{r},t) = \alpha_{WB}^{(l)}(s,t,\sigma_n,\mathbf{v}_b)WB_{tf}^{(l)}(\mathbf{r}-\mathbf{v}_b,t-1) + (1-\alpha_{WB}^{(l)}(s,t,\sigma_n,\mathbf{v}_b))WB_n^{(l)}(\mathbf{r},t) \tag{10}$$

where $WB_{tf}^{(l)}$ stands for the temporally processed wavelet band at scale l. The weighting factor $\alpha_{WB}^{(l)}(s,t,\sigma_n,\mathbf{v}_b)$ controls the amount of filtering for each wavelet band ($WB^{(l)}$) in the following way:

$$\alpha_{WB}^{(l)}(s,t,\sigma_n,\mathbf{v}_b) = b_{WB}^{(l)}(\vartheta_{WB}^{(l)}(s,t,\sigma_n,\mathbf{v}_b))^2 \tag{11}$$

with $\vartheta_{WB}^{(l)}$ the motion *reliability* per wavelet band WB of the motion vector $\mathbf{v}_b$ and $b_{WB}^{(l)}$ a normalizing parameter that we experimentally optimize in terms of the mean squared error. Specifically, in our two-scale decomposition implementation we determine the parameters for the first (finest) scale as $b_{WB}^{(1)} = 0.9$, for the second scale as $b_{WB}^{(2)} = 0.95$ and $b_{LL}^{(2)} = 1.25$ for the low-pass (approximation) band (roughest scale). We have chosen the quadratic dependency in (11) because our experiments showed that it introduces less temporal blur and artifacts than e.g., a linear model which does not respond well to motion vector miss-matches. In addition, we have also tested higher degree models and found that the quadratic dependency model indeed provides the best results in terms of maximal PSNR of the denoised sequence.

Because of the imperfections of the motion estimation process (due to various difficulties such as occlusion or an imperfect motion estimation model) and rather rough assumption that the noise is white Gaussian, the temporal filter still leaves some noise behind.

2.2. 2D spatially adaptive speckle suppression scheme

In this section we propose a denoising algorithm of [8] for removing the remaining noise. The adaptive speckle suppression scheme from of [8] is a shrinkage technique where each denoised wavelet coefficient $WB_{stf}^{(l)}(\mathbf{r})$ is determined as $WB_{stf}^{(l)}(\mathbf{r}) = q(\mathbf{r})WB_{tf}^{(l)}(\mathbf{r})$, where q represents shrinkage parameter. Specifically, q reduces wavelet coefficient amplitude depending on the estimated probability that it represents important noise-free image feature. The corresponding estimated probability is computed based on two sources of information: (i) coefficient magnitude and the corresponding conditional probability density functions of the coefficient magnitudes representing mainly noise on the one hand, and representing the useful edges on the other hand and (ii) spatial context information.

Let us denote the coefficient magnitude by $m(\mathbf{r})$, and let us associate with each wavelet band $WB_{tf}^{(l)}$ a mask $\mathbf{X}$ of binary labels, where $X(\mathbf{r}) = 0$ ("non-edge" label) if $WB_{tf}(\mathbf{r})$ represents mainly noise, and $X(\mathbf{r}) = 1$ ("edge" label) if $WB_{tf}(\mathbf{r})$ represents useful signal. The likelihood of $m(\mathbf{r})$ given the label value $X(\mathbf{r})$ will be denoted by $p(m(\mathbf{r})|X(\mathbf{r}))$.

The method of [8] defines the shrinkage parameter as follows:

$$q(\mathbf{r}) = P(X(\mathbf{r}) = 1|m(\mathbf{r}),\mathbf{X}_{\mathbf{r}'}) \tag{12}$$

where $\mathbf{X}_{\mathbf{r}'}$ denotes the mask labels at all positions except $\mathbf{r}$. It is shown in [8] that this expression can be expressed as:

$$q(\mathbf{r}) = \frac{\varepsilon(\mathbf{r})\eta(\mathbf{r})}{1+\varepsilon(\mathbf{r})\eta(\mathbf{r})} \tag{13}$$

where $\varepsilon(\mathbf{r})$ is the likelihood ratio at the current position $\mathbf{r}$:

$$\eta(\mathbf{r}) = \frac{p(m(\mathbf{r})|X(\mathbf{r})=1)}{p(m(\mathbf{r})|X(\mathbf{r})=0)} \tag{14}$$

and $\eta(\mathbf{r})$ is derived from a spatial surrounding $\partial(\mathbf{r})$ as follows [8]:

$$\eta(\mathbf{r}) = \exp(\gamma \sum_{\mathbf{r}_k \in \partial(\mathbf{r})} (2X(\mathbf{r}_k) - 1)) \tag{15}$$

Note that if there is equal number of "edge" and "non-edge" labels in the neighborhood $\partial(\mathbf{r})$ then the neighborhood information is "neutral" and (13) yields $\eta(\mathbf{r}) = 1$. This means that in case where there is equal number of edge and non-edge labels in $\partial(\mathbf{r})$ the spatial context information does not influence the calculation of the shrinkage factor in (12). In that case, the suppression factor depends on the coefficient magnitude and the corresponding likelihood ratio only. If the neighborhood $\partial(\mathbf{r})$ contains a majority of edge ("1") labels then $\eta(\mathbf{r})$ is greater than 1 and it increases the shrinkage factor in (12). The opposite is true when there is a majority of "non-edge" (zero) labels in the corresponding neighborhood. In this way, the spatial context information, expressed through $\eta(\mathbf{r})$, favors suppressing less those wavelet coefficients that are surrounded by the majority of "edge" coefficients and it favors suppressing heavier the isolated big coefficients (that are more likely to originate from noise).

The binary labels $\mathbf{X}$ are estimated using a preliminary coefficient classification as follows:

$$X(\mathbf{r}) = \begin{cases} 0, & |WB_{tf}^{(l)}(\mathbf{r})||WB_{tf}^{(l+1)}(\mathbf{r})| < \sigma_l^2 \\ 1, & |WB_{tf}^{(l)}(\mathbf{r})||WB_{tf}^{(l+1)}(\mathbf{r})| \geq \sigma_l^2 \end{cases} \tag{16}$$

where σ_l is the standard deviation of noise at the scale l. In this process, a "coarse-to fine" strategy is employed, where a previously denoised wavelet subband from a coarser scale is used to estimate the edge positions at the next finer scale.

For the calculation of the likelihood ratio's, the following model was proposed in [8]:

$$\begin{aligned} p(m(\mathbf{r})|X(\mathbf{r}) = 0) &= \frac{1}{a} \exp(-m(\mathbf{r})/a) \\ p(m(\mathbf{r})|X(\mathbf{r}) = 1) &= \frac{m(\mathbf{r})}{2b^3} \exp(-m(\mathbf{r})/b) \end{aligned} \tag{17}$$

where the parameters a and b are computed automatically during the denoising procedure, by minimizing the mean squared difference between the normalized histogram (of data for binary labels 0 and 1) and the corresponding sampled analytic approximations given in (17), i.e., by maximum likelihood approach.

2.3. The proposed 3D extension

In this work, we propose a 3D extension of this method, by modifying the equation 15 in such a way that we do not take only mask from the current frame into account but from the previous frame as well, in case when no undergoing motion is detected (assumed) between those frames. Specifically, the considered spatial surrounding ∂ is now extended to the 3D (2x3x3) surrounding, where the binary labels from masks are considered from both the current and previously processed frame. The criteria for deciding whether only 2D surrounding will be considered or the 3D one is defined as follows: If

$$||WB_{tf}^{(l)}(r,t)WB_{tf}^{(l+1)}(r,t)| - |WB_{tf}^{(l)}(r,t-1)WB_{tf}^{(l+1)}(r,t-1)|| > \sigma_l^2 \tag{18}$$

is satisfied then only 2D neighbouring surrounding is used as suggested in [8]; otherwise 3D (2x3x3) surrounding is used for computing $\eta(r)$ in equation 15

3. EXPERIMENTAL RESULTS

The performance of the proposed algorithm is evaluated on the processed 3D OCT images of "phantom" and human tissue, provided by AGFA. Images denoised using different methods were compared using contrast-to-noise ratio (CNR) and the equivalent number of looks (ENL) over six regions of interest in each image. CNR is a measure of contrast between image feature and backgound of image, while the ENL measures smoothness in areas that should be homogeneous but are corrupted by speckle. Contrast-to-noise ratio is defined as $CNR_m = 10log[(\eta_m - \eta_b)/\sqrt{(\sigma_m^2 + \sigma_b^2)}]$ and $ENL_m = \eta_m^2/\sigma_m^2$, where η_m and σ_m are the mean and the standard deviation of the pixel values in the m-th region of interest (ROI) and η_b and σ_b are the mean and standard deviation of the background region. Besides these measures, denoised images are also compared using global image sharpness parameter, β which is used to determine the degree of smoothing of denoised images, and is defined [14]. Our method is compared to the methods of Adler and Fujimoto [15], Lee [16], Balster [17] and rotating kernel method presented in [18]. Some of the methods used for comparison were developed specifically for OCT images ([15], [18]), while others like Balster and were proposed only for white Gaussian noise. Nevertheless we used them here for comparison of speckle noise in OCT images. Regions of interest are chosen in a way that characteristic textures from the image are present. First region represents background of the image. results of quantitative comparisons of the proposed and reference methods are given in the tables 1 and 2

Table 1. Image quality metrics for the compared methods for breast cancer image

Method	CNR	ENL
"RKT"	3.75	1843
"Lee"	4.65	2756
"Balster"	4.23	2430
"Adler"	3.17	1289
"Proposed"	5.04	2889

Table 2. Image quality metrics for the compared methods for phantom image

Method	CNR	ENL
"RKT"	3.15	1102
"Lee"	3.29	1205
"Balster"	2.38	639
"Adler"	2.33	574
"Proposed"	3.63	1933

As can be seen from the tables, proposed method outperforms reference methods in terms of objective measures. Proposed method also outperforms reference methods in the terms of visual quality. Denoised images of noisy fantom image and the image of breast cancer are shown in Figures 3 and 2

As can be seen from 2 and 3 methods of [17] and [16] remove background noise better than other reviewed methods. Still details are slightly better preserved in image processed by algorithm of [16]. The largest amount of the backgound noise remains after noise removal using algoritm of [18]. Method of [15] removes background noise very well, while preserving details better than the other reference methods. Proposed method outperforms reference methods in terms of visual quality, both in backgound noise removal and detail preservation. This can especially be observed for image of human tissue.

4. CONCLUSION AND FUTURE WORK

In this paper we have proposed a novel scheme for removing noise in 3D OCT images. In the proposed scheme we first apply the temporal filter of [7] and subsequently remove the remaining noise by a simple 3D extension of the despeckling filter from [8].

In future we intend to model, in a more sophisticated manner, the statistical and correlation properties of the noise in OCT images and use it in both temporal and spatial filter for optimal filtering performance.

REFERENCES

1. B.E. Bouma and G. J. Tearney, *Handbook of Optical Coherence Tomography*, Marcel Dekker, Inc., New York, 2002.
2. S. Bourquin, A.D. Aguirre, I. Hartl, P. Hsuing, T.H. Ko, and J.G. Fujimoto, "Ultrahigh resoltion real time oct imaging using a compact femtosecond nd:glass laser and nonlinear fiber," *Optics Express*, vol. 11, no. 24, pp. 3290–3297, 2003.
3. R. Bernstein, "Adaptive nonlinear filters for simultaneous removal of different kinds of noise in images," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 1275–1291, 1987.
4. G. Franceschetti, V. Pascazio, and G. Shirinzi, "Iterative homomorphic technique for speckle reduction in synthetic-aperature radar imaging," *Journal of Optical Society of America*, vol. 12, pp. 686–694, 1995.
5. S.H. Xiang, L. Zhou, and J.M. Schmitt, "Speckle noise reduction for optical coherence tomography," *Proc. SPIE (Euroser)*, vol. 3196, pp. 79–88, 1997.
6. D.T. Kuan, A.A. Sawchuk, T.C. Strand, and P. Chavel, "Adaptive noise smoothing for images with signal-dependent noise," *IEEE Trans. Pattern Anal. MAch. Intell.*, vol. 7, pp. 165–177, 1985.
7. V. Zlokolica, A. Pizurica, and W. Philips, "Wavelet-domain video denoising based on reliability measures," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 8, no. 16, pp. 993–1007, Aug. 2000.
8. A. Pizurica, A. Philips, I. Lemahieu, and M. Acheroy, "Despeckling sar images using wavelets and a new class of adaptive shrinkage estimators," in *IEEE International Conference on Image Processing*, Thessaloniki, Greece, Sept. 2001, pp. 233–236.
9. S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, July 1992.
10. V. Zlokolica, A. Pizurica, and W. Philips, "Wavelet domain noise-robust motion estimation and noise estimation for video denoising," in *First International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, Scotssdale, Arizona, USA, 2005.
11. G. De Haan, *Video Processing for multimedia systems*, University Press Eindhoven, Eindhoven, The Nederlands, 2003.
12. A.M. Tekalp, *Digital Video Processing*, Prentice Hall PTR, Upper Saddle River, New Jersey, 1995.
13. R. Li and M.L. Lio, "A new three-step search algorithm for block motion estimation," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 4, no. 4, pp. 438–442, Aug. 1994.
14. F. Sattar, L. Floreby, G. Salomonsson, and Lovstrom B., "Image enhancement based on a nolinear multiscale method," *IEEE Trans. on Image Processing*, vol. 6, no. 6, pp. 888–895, 1997.
15. D. C. Adler, T. Ko, and Fujimoto J., "Speckle reduction in optical coherence tomography images by use of a spatially adaptive wavelet filter," *Optics Letters*, vol. 29, no. 24, pp. 2878–2880, 2004.
16. J.S. Lee and K. Hoppel, "Noise modeling and estimation of remotely-sensed images," in *International Geoscience and Remote Sensing*, Vancouver, Canada, 2001.
17. E. J. Balster, Y. F. Zheng, and R. L. Ewing, "Feature-based wavelet shrinkage algorithm for image denoising," *IEEE Trans. on Image Processing*, vol. 14, no. 12, pp. 2024 – 2039, Dec. 2005.
18. Rogowska J. and Brezinski M., "Evaluation of the adaptive speckle supression filter for coronary optical coherence tomography imaging," *Optics Letters*, vol. 29, no. 24, pp. 2878–2880, 2004.

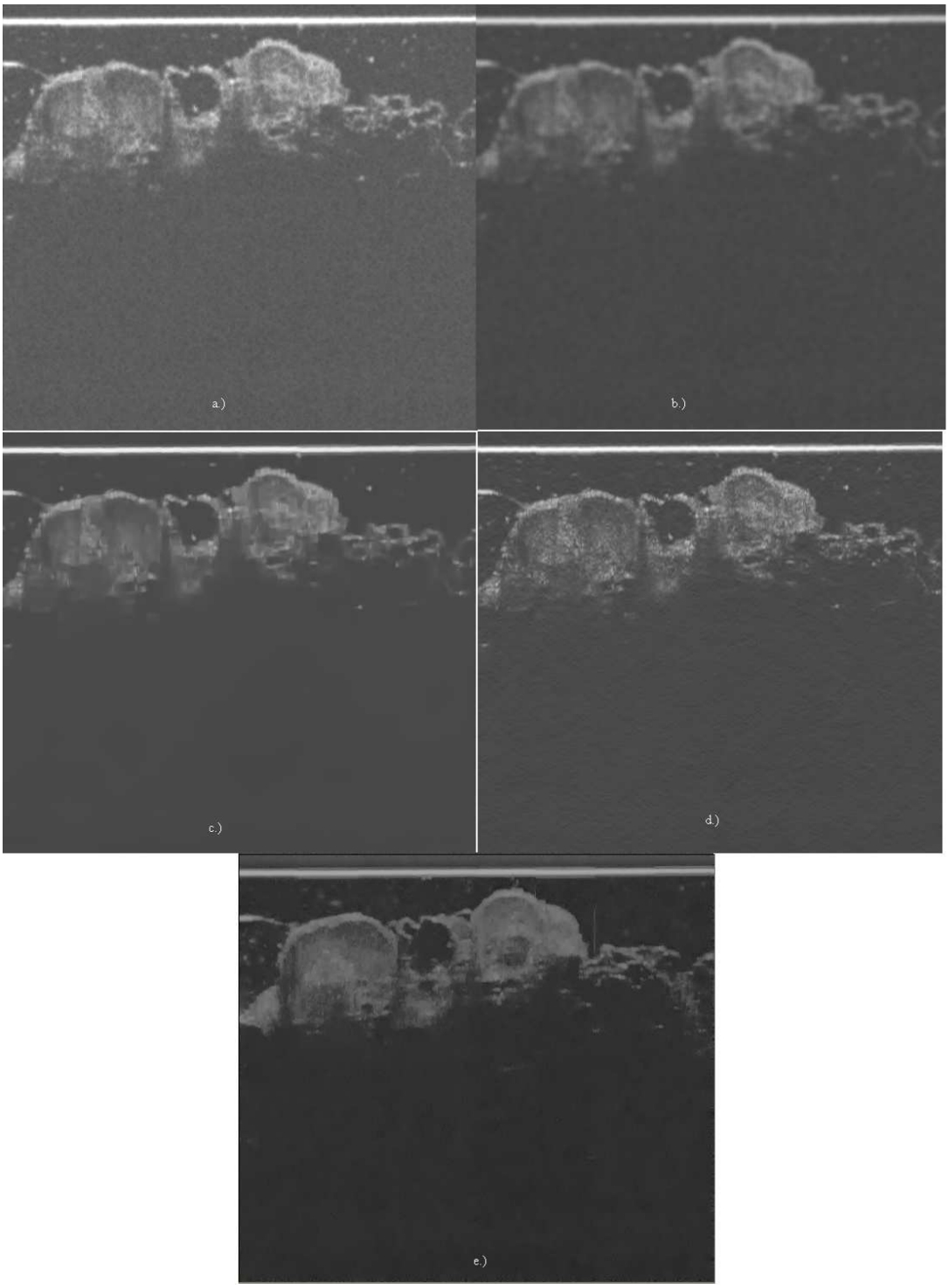

Figure 2. OCT image of the breast cancer denoised using a.) Rotating kernel transform method of [18] b.) Method of Lee [16] c.) Method of Balster [17] d.) Method of Adler [15] e.) Proposed method

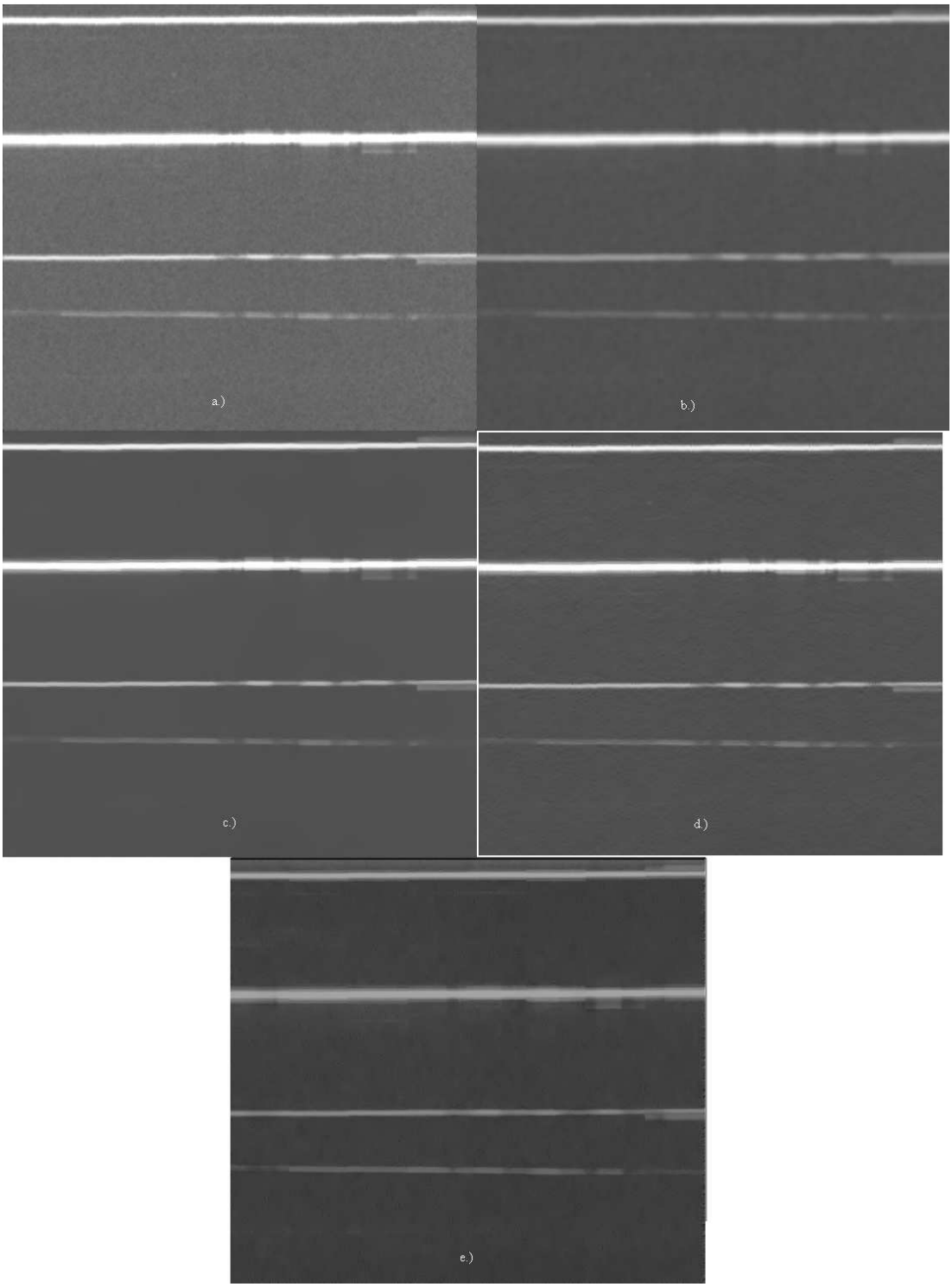

Figure 3. OCT "phantom" image denoised using a.) Rotating kernel transform method of [18] b.) Method of Lee [16] c.) Method of Balster [17] d.) Method of Adler [15] e.) Proposed method